



Published in final edited form as:

Neuropsychology. 2015 November ; 29(6): 940–948. doi:10.1037/neu0000208.

Absence of Practice Effects in Preclinical Alzheimer's Disease

Jason Hassenstab, Ph.D.^{1,2}, David Ruvolo, M.S.¹, Mateusz Jasielc, M.S.³, Chengjie Xiong, Ph.D.³, Elizabeth Grant, Ph.D.³, and John C. Morris, M.D., FAAN^{1,4,5,6}

David Ruvolo: ruvalod@abraxas.wustl.edu; Mateusz Jasielc: Mateusz@wubios.wustl.edu; Chengjie Xiong: chengjie@wubios.wustl.edu; Elizabeth Grant: betsy@wubios.wustl.edu; John C. Morris: morrisj@abraxas.wustl.edu

¹Charles F. and Joanne Knight Alzheimer's Disease Research Center, Department of Neurology, Washington University School of Medicine, Washington University in St. Louis, St. Louis, MO. USA

²Department of Psychology, Washington University in St. Louis, St. Louis, MO. USA

³Division of Biostatistics, Washington University in St. Louis, St. Louis, MO. USA

⁴Department of Pathology and Immunology, Washington University in St. Louis, St. Louis, MO. USA

⁵Department of Physical Therapy, Washington University in St. Louis, St. Louis, MO. USA

⁶Department of Occupational Therapy, Washington University in St. Louis, St. Louis, MO. USA

Abstract

Objective—To describe how practice effects influence cognitive trajectories and determine if a reduction in practice effects is a potential marker of Stage 3 preclinical AD.

Method—Participants included 263 older adults who were cognitively normal at baseline (CDR=0) and returned for an average of 9.5 annual visits. Participants completed standard tests of episodic memory, visuospatial ability, semantic memory, and executive function. Progressors (n=66) converted to CDR>0 with a diagnosis of symptomatic AD after a minimum of 3 visits and Stable participants (n=197) never progressed to CDR>0. Practice effects, defined as the slope of performance across visits 1–3, were compared between groups and used within-subjects to predict risk of conversion. Change-point models that account for retest were contrasted with linear models that ignore retest.

Results—The Stable group showed practice effects on episodic memory measures ($\beta = 0.14$, SE = .02, $p < .0001$) but the Progressor group did not ($\beta = 0.03$, SE = .03, $p = 0.343$). Across all participants, practice effects on episodic memory tests were associated with a decreased risk of progression to AD (SHR = 0.110, 95% CI 0.032–0.384, $p = .001$). Finally, use of change-point models dramatically altered rate of change estimates compared with models that ignored practice.

CORRESPONDENCE: Jason Hassenstab, Ph.D., 4488 Forest Park Ave., Suite 130, St. Louis, MO 63108. Office: 314-747-4032, Fax: 314-286-2443, hassenstabj@wustl.edu.

AUTHOR CONTRIBUTIONS:

Study concept and design: Hassenstab, Ruvolo, Morris

Data acquisition: Grant, Ruvolo

Data analysis and interpretation: Hassenstab, Ruvolo, Morris, Xiong, Jasielc

Drafting/revising manuscript for content: All authors.

Conclusions—Our results indicate that preclinical AD is marked by a reduction in practice effects in episodic memory and that the magnitude of gain from retesting is inversely related to progression risk. Assessment of practice effects may be a face-valid indicator of Stage 3 preclinical AD.

Keywords

Alzheimer's disease; dementia; practice effects; memory; retest

Introduction

In studies of aging and dementia, repeated neuropsychological assessment at regular intervals is critical for detecting cognitive change and monitoring dementia progression. However, cognitive performance typically improves with retesting as participants become familiar with the test items, the testing environment, and test concepts and strategies. Such improvement can complicate assessment of cognitive change (Lievens, Buyse, & Sackett, 2005). The magnitude of performance gains over time, termed *practice effects*, appears to be influenced by age, educational attainment, IQ, disease status, and ability levels in specific domains of cognition such as procedural memory (McCaffrey, Duff, & Westervelt, 2000). Meta-analytic studies have found that episodic memory abilities best predict the magnitude of Practice effects (Lievens, Reeve, & Heggstad, 2007) with typical performance gains in healthy adults of approximately one quarter of a standard deviation from the initial administration to the first retest (Calamia, Markon, & Tranel, 2012; Hausknecht, Halpert, Di Paolo, & Moriarty Gerrard, 2007). The first two retests demonstrate the greatest practice effects, which then diminish with further retesting (Ivnik et al., 2000; Thorndike, 1922). A spurious accelerated decline or inflection point may occur after the third assessment that may not be related to disease progression but instead represents the declining influence of practice effects (Dodge, Wang, Chang, & Ganguli, 2011). Thus, practice effects must be factored into the interpretation of cognitive trajectories. In progressive dementing disorders such as Alzheimer's disease (AD), a lack of performance gains upon retesting can be interpreted either as stable cognitive performance or as an absence of practice effects due to underlying disease. Others have shown that an absence of practice effects in clinically symptomatic individuals with amnesic Mild Cognitive Impairment (aMCI) due to AD is associated with cognitive decline and worse clinical outcome (Cooper, Lacritz, Weiner, Rosenberg, & Cullum, 2004; Duff et al., 2007; 2011).

There is now ample evidence from clinicopathological and molecular biomarker studies that the pathophysiological process of AD may begin more than a decade prior to the onset of clinically overt cognitive and functional decline (Bateman et al., 2012; Price & Morris, 1999). However, preclinical AD may not be benign as subclinical cognitive deficits have been reported in cognitively normal (CN) individuals who at autopsy have neuropathological AD (Price et al., 2009) and the absence of practice effects in CN persons has been shown to predict autopsy-confirmed AD dementia (Galvin et al., 2005). A hypothetical model of stages of preclinical AD proposes that the stage just prior to the appearance of clinically evident cognitive deficits is marked by "subtle cognitive decline" (Sperling et al., 2011) and, although these subtle changes have been variously

operationalized, there now is emerging support that their presence is highly predictive of progression to symptomatic AD (Knopman et al., 2012; Vos et al., 2013). To build on the relatively few studies of practice effects in asymptomatic individuals who later develop cognitive impairment (Howieson et al., 2008; Machulda et al., 2013), we aimed to explore whether the “subtle cognitive decline” proposed for Stage 3 of preclinical AD includes the absence of practice effects. We hypothesized that CN participants who later develop clinically manifest cognitive impairment would show less practice effects compared with CN participants who remain stable. We also hypothesized that the magnitude of performance gains across the first three assessments would predict progression to symptomatic AD. Finally, we hypothesized that ignoring the influence of practice effects on cognitive trajectories would obscure estimates of cognitive decline.

Methods

Participants

The study population was drawn from participants enrolled in longitudinal studies of aging and dementia at the Knight Alzheimer’s Disease Research Center (ADRC). Longitudinal clinical and cognitive data collected from December of 1979 through December of 2012 at the Knight ADRC were used for analysis. All participants provided written informed consent and the studies were reviewed and approved by the Human Research Protections Office at Washington University in St. Louis.

Inclusion/Exclusion criteria

Participants were cognitively normal at study entry (baseline) according to a CDR rating of 0, remained CDR 0 for at least two additional annual assessments after baseline, completed at least four consecutive annual assessments, completed neuropsychological testing at each annual assessment, and were aged 65 years and older. Participants were excluded for incomplete cognitive or other clinical data, if their primary clinical diagnosis after visit three was uncertain dementia or dementia of an etiology other than AD (i.e., Parkinson’s or Lewy Body dementia), or if their CDR rating was inconsistent at any timepoint (i.e., fluctuating between CDR=0 and CDR=0.5). It is important to note that we selected participants who were clinically normal according to the CDR for at least three consecutive assessments, and that during the first three assessment periods these participants would not be considered symptomatic by established criteria for mild cognitive impairment (MCI; Albert et al., 2011; Petersen et al., 1999). In this study, for initially CN individuals who developed CDR 0.5, only those with an etiological diagnosis of AD were included. Many of our CDR 0.5/AD individuals elsewhere are considered as MCI (Storandt, Grant, Miller, & Morris, 2006), or more recently as “MCI due to AD” (Albert et al., 2011). We use the term “symptomatic AD” to encompass all individuals from MCI to overt dementia when the etiological diagnosis is AD, which is confirmed at autopsy in 93% of cases (Berg et al., 1998).

Participant Classification

Progressors (n=66) progressed to a CDR=0.5 with a diagnosis of symptomatic AD after their third annual visit and remained CDR>0 for at least one additional annual visit. Stable participants (n=197) remained CDR=0 during follow-up.

Practice Effects Definition

The goal of this study was to define practice effects as they are observed in a longitudinal cohort study, where participants are seen at regular (typically yearly) intervals. Shorter retest intervals produce the largest practice effects (Theisen, Rapport, Axelrod, & Brines, 1998), but practice effects remain evident with as much as 7 years between assessments (Salthouse, Schroeder, & Ferrer, 2004). Practice effects are typically most pronounced between the first and second retest, with a diminishing effect on subsequent assessments (Calamia et al., 2012; Ivnik et al., 2000; Thorndike, 1922). Thus, we chose to model the slope of cognitive performance across the first three yearly assessments and prior to clinical diagnosis in order to best describe practice effects in the preclinical stage of AD. Sampling practice effects over longer periods of retest may further compromise our ability to parse practice effects from age and disease related cognitive decline (Ivnik et al., 2000).

Clinical assessment

The Clinical Dementia Rating scale (CDR(Morris, 1993) was administered by a CDR certified physician or nurse-clinician at each annual clinical assessment. The CDR is derived from semi-structured interviews conducted separately with both the participant and a knowledgeable informant. All participants complete a comprehensive health history, depression screen, aphasia battery, and neurological examination at each annual assessment. The CDR is determined independently of the neuropsychological assessment.

Neuropsychological assessment

All participants completed neuropsychological assessment within approximately 2 months of their clinical evaluation by trained psychometricians. Standardized test scores were averaged to form four composites derived from factor analytic studies (Johnson, Storandt, Morris, & Galvin, 2009; Johnson, Storandt, Morris, Langford, & Galvin, 2008). The episodic memory composite included the Free and Cued Selective Reminding Test free recall score (FCSRT(Grober, Buschke, Crystal, Bang, & Dresner, 1988), Associate Learning from the Wechsler Memory Scale (WMS(Wechsler & Stone, 1973) and the immediate and delayed Recall from the WMS-R Logical Memory test (Wechsler, 1987). The semantic composite included the Information subtest from the Wechsler Adult Intelligence Scale (WAIS(Wechsler, 1955), the Boston Naming Test (Goodglass & Kaplan, 1983), and Animal Naming (Goodglass & Kaplan, 1983). The executive function composite included WMS Mental Control, Digit Span, and Letter Fluency for S and P. The visuospatial composite included the WAIS Block Design and Digit Symbol subtests and Trailmaking Test A and B (Armitage, 1946). A global cognitive composite was created from the average of the four domain-specific composites. The reference (normative) group used to standardized the tests prior to forming the composites was a sample of 310 people from the Knight ADRC (age = 74.5 +/- 8.6 years; education = 14.8 +/- 3.2 years) who were enrolled as CDR=0, had at least one annual follow-up, and never progressed to CDR>0 (Johnson et al., 2009).

Statistical analysis

Demographic characteristics were compared with t-tests for quantitative variables and with Chi-Square tests for nominal variables. Analysis of covariance was used to test for

differences in baseline cognitive performance, with age, gender, and education as covariates. Linear mixed effects models examined the interactive effects of group (Progressor vs. Stable) and time (in years from baseline) on cognitive performance across the first three annual assessments using lme4 (Bates, 2010) in the R statistical software package (www.r-project.org). Covariates included fixed effects for baseline cognitive score, age at baseline, the interaction between baseline age and time (years from baseline), gender, education, presence of one or more apolipoprotein ϵ 4 allele (APOE4), time, and group. Random effects were specified for participants and time. Effects from APOE4 were not significant in any models and were not included in final analyses. A competing-risks survival analysis using Fine and Gray's subdistribution hazards model (SHR; (Fine & Gray, 1999), with time on study as the timescale, was used to determine if practice effects slopes were associated with risk of progression to CDR>0. While both SHR and standard Cox hazards models treat dropout as censored, SHR differs in that mortality is not censored but regarded as a competing event that can impede progression to symptomatic AD. The PE slopes for each individual were computed using ordinary least squares regression. SHRs were implemented using the STCRREG command in STATA v.12 (StataCorp. College Station, TX). Covariates included baseline cognitive performance, age at baseline, gender, education, and APOE4 status. To assess the sensitivity of the survival analyses results to the choice of time scale and model specification related to age at baseline, analyses were repeated with age as the time scale accounting for left truncation, including and excluding age at baseline as a covariate, as well as stratifying by baseline age quartiles. These three alternative approaches were similar to results using time-on-study as the time scale. Finally, performance trajectories in all participants were modeled with linear mixed effects models, first assuming a single linear slope over all assessments and again with a spline model with a changepoint at the third assessment. This modeled the effects of practice separately from the slope of performance over all following annual assessments.

Results

Table 1 provides demographic characteristics and baseline cognitive data. Participants were mostly Caucasian (90.1%). The Stable group was younger at baseline (71.8 years) and had more females (60.4% female) than Progressors (73.5 years and 54.6% female). Most participants were college educated (Progressors=15.3 years vs. Stable=15.6 years). Progressors had more visits (10.1 +/- 4.1 assessments) than stable participants (6.6 +/- 3.9 assessments). Baseline comparisons of cognitive composites, controlling for age, gender, and education, revealed that Progressors performed worse than the Stable group on the global cognitive composite ($t(251) = 2.00, p = .047$) and the visuospatial composite ($t(251) = 2.15, p = .032$). Comparison of performance on individual cognitive measures at baseline revealed worse performance in the Progressors on Logical Memory Immediate recall ($t(139) = 2.62, p = .009$) and the Trailmaking Test Part A ($t(256) = -2.55, p = .011$). See Table 2 for cognitive performance across visits 1–3.

Analyses of practice effects slopes (Table 3) revealed significant group by time (defined as years since baseline) interactions on the episodic memory composite, wherein the Stable group exhibited more performance gains ($\beta = 0.14, SE = .02, p < 0.001$) than the Progressors ($\beta = 0.03, SE = .03, p = 0.343$; Figure 1). As mentioned above, the episodic memory

composite is composed of three separate episodic memory tasks: Logical Memory, Associate Memory, and the FCSRT. To determine if one or more individual tests were driving the effect of the composite result, each test was modeled individually using z-scores. Only Logical Memory delayed recall exhibited a group by time interaction, with the Stable group showing a performance gain across the first three assessments ($\beta = 0.14$, $SE = .03$, $p < 0.001$) compared to the Progressors ($\beta = -0.16$, $SE = .10$, $p = 0.138$). There were no interactive effects of group and time on the visuospatial, semantic memory, executive function, or global cognitive composites.

In the competing risks survival analyses, the slopes of the cognitive composites across the first three assessments were used to predict risk of CDR progression. Holding baseline performance, age, gender, education, and APOE4 status constant, a one unit increase in the estimated annual slope from baseline to the third assessment in episodic memory performance was associated with an 89.0% reduced risk of CDR progression (SHR = 0.11, 95% CI 0.032–0.384; $p = .001$; see Figure 2). Slopes of the visuospatial, executive function, semantic, and the global composites were unrelated to risk of progression.

To determine if accounting for practice effects provided a better estimate of cognitive change, we ran two models on all our participants from baseline through their last available timepoint using the composites for episodic memory and global cognition. The first model ignored practice effects and assumed a single slope of cognitive performance across time. The single slope indicated no changes in performance over time on episodic memory ($\beta = .005$, $SE = .011$, $p = .656$) but there was a decline on the global composite ($\beta = -.0141$, $SE = .006$, $p = .020$). The second model considered rates of changes from the first three annual assessments (which putatively capture the period of maximum practice effects) separately from performance on all following annual assessments (Table 4). There were positive slopes across the first three timepoints on both episodic memory ($\beta = .117$, $SE = .017$, $p < .001$) and the global composite ($\beta = .046$, $SE = .008$, $p < .001$), indicating an effect of practice. The slope of cognitive performance in remaining assessments was negative for both the episodic memory composite ($\beta = -.0640$, $SE = .013$, $p < .001$), and the global composite ($\beta = -.0483$, $SE = .008$, $p < .001$). Overall, these results indicate that episodic memory and global cognition exhibit practice effects and that failure to consider practice effects in analyses could mask cognitive decline.

Discussion

The goals of this study were to describe how practice effects influence cognitive functioning in a longitudinal cohort study of older adults and to determine if an attenuation of practice effects during the asymptomatic preclinical phase of AD would predict progression to symptomatic AD. Overall, our results suggest that practice effects dramatically impact cognitive trajectories in CN older adults and may also distort the true pace of cognitive decline in the preclinical stage of AD. Asymptomatic participants who later progressed to symptomatic AD showed greatly reduced performance gains on episodic memory tasks compared to participants who remained cognitively stable, indicating that a reduction in practice effects observed at yearly intervals may be a cognitive predictor of incident symptomatic AD. We also found that the degree of attenuation of practice effects in episodic

memory performance was inversely associated with progression risk. Participants who exhibited year-to-year improvement on episodic memory performance had a substantially reduced risk of progressing to symptomatic AD, suggesting that a reduced ability to learn from prior assessments may be an indicator of Stage 3 preclinical AD. Finally, we demonstrated that ignoring practice effects on tests of episodic memory could obscure estimates of cognitive change and lead to misinterpretation of results.

Of all of the cognitive composites, including executive function, semantic, visuospatial, and the global composite, only episodic memory exhibited a significant difference in practice effects between the Progressors and the Stable group. Subtle changes in episodic memory are consistent with expectations in the context of preclinical AD progression. According to the staging theories of preclinical AD (Jack et al., 2010; 2013; Sperling et al., 2011), and subsequently proven in studies of autosomal-dominant mutation carriers (Bateman et al., 2012), subtle cognitive changes begin in Stage 3, the penultimate stage prior to clinical diagnoses such as MCI or very mild AD (CDR=0.5 AD). It is well known that episodic memory is associated with the structure and function of the medial temporal lobes including the hippocampus (Aggleton & Brown, 1999; Tulving, 2002) as well as the network of cortical regions known as the default mode network (Buckner et al., 2005). Several studies have shown that changes in these regions may begin prior to the appearance of clinical symptoms of AD (Dickerson et al., 2011; Price & Morris, 1999; Sheline et al., 2010). Without consideration of practice effects, a lack of intraindividual declines in episodic memory in the years immediately preceding diagnosis could be considered evidence of stable cognitive performance.

We next sought to determine if the slope of practice effects on episodic memory could be used to predict progression to symptomatic AD. We found that the magnitude of attenuation of practice effects on episodic memory tasks was inversely related to the risk of progression to symptomatic AD. Participants exhibiting significant practice effects had an 89% reduced risk of progression. Taken together, results from mixed model analyses combined with results from survival analyses represent an important finding: a lack of improvement upon retesting, rather than decline, might represent a clinically relevant manifestation of Stage 3 preclinical AD. These results also highlight the importance of tracking intraindividual change in cognitive status with serial neuropsychological assessment.

Significant practice effects were only observed on tests of episodic memory, but there were less consistent group differences in baseline performance on individual episodic memory measures, which suggests that our results may represent a more subtle deficit in memory systems that may not be detected by relying on mean score differences in cross-sectional analyses. It is important to consider what an attenuation of practice effects represents and whether it indicates a broader phenomenon than subtle impairment in episodic memory. The literature suggests that practice effects can be observed across a wide range of retest intervals (Salthouse et al., 2004) and that episodic memory abilities are commonly affected in the preclinical stage of AD (Grober et al., 2008; Knopman et al., 2012; Vos et al., 2013). Although the Progressors were beginning to show cognitive decline, they were still cognitively normal according to the CDR. It is likely that some information was retained from the prior visit, even after a one-year retest interval, and perhaps more so on measures

where test instructions direct participants to store information for later recall, as was the case with our associative recall and narrative recall measures. By design, list learning measures provide increased exposure to test items during the learning phase and encourage a particular recall strategy that may carryover in subsequent assessments (Grober et al., 1988; Tulving, 1967). This may be particularly evident on memory tests with a delayed recall condition where exposure and strategy use are further reinforced. The fact that we observed practice effects only on episodic memory measures may be in part due to the increased exposure to the material and the emphasis on test-taking strategies that is implied by the test instructions. Thus, there is an aspect of procedural memory as well as item-specific memory that may carryover from prior administrations. Procedural memory, especially for motor-related activities, is thought to be preserved in early AD (Zanetti et al., 2001), but it may be the case that more subtle forms of procedural memory such as test taking strategies show subtle decline in preclinical disease.

Failure to account for practice effects in serial testing can result in spurious interpretations. This risk was made clear in analyses in which episodic memory slopes across all participants were first modeled linearly with no assumption of the influence of practice effects and again with a changepoint placed at the third assessment (Table 4). Ignoring practice effects produced a flat slope suggestive of no memory decline over a mean follow-up period of 9.5 visits (11.1 years). The addition of the changepoint yielded dramatically different results. Slopes after the third assessment exhibited highly significant declines on episodic memory, suggesting that episodic memory shows strong practice effects and that these effects may 'mask' cognitive changes due to aging and disease processes. The risks of ignoring practice effects have relevance for cognitive enhancement trials as well. Measures that exhibit large practice effects may make it difficult to determine whether an intervention has truly evidenced a clinically meaningful improvement in cognitive functioning (Goldberg, Keefe, Goldman, Robinson, & Harvey, 2010).

To accurately estimate cognitive change, it may be critical to model practice effects relative to the number of prior study assessments. For example, it is common in analysis of registry datasets to stratify by age ranges or other criteria without consideration of the number of prior test administrations. This poses a risk of inaccurately estimating rate of change in cognition. Participants may exhibit a different slope during their initial assessments due to practice effects than during follow-up assessments when practice effects have less influence, thus confounding estimates of cognitive change. One possible solution would be to model a practice effect with a diminishing return that accounts for the number of prior assessments.

Another consideration is measure-specific practice effects, independent of the number of prior assessments. This may be important when a new measure is added to the cognitive battery of an ongoing study. In this scenario, some aspects of practice effects such as testing environment familiarity and testing expectations will be diminished. However, participants with several years of prior visits will likely show practice effects on a new measure added to a study protocol, especially a measure of episodic memory. Future studies should carefully consider modeling practice effects based on number of prior test administrations.

There are several limitations in our study. Our measures of cognition are widely used in aging and dementia studies but due to the historical legacy of our dataset are somewhat limited in their sensitivity and in scope of cognitive domains. We do not, for example, include sensitive measures of attentional control, processing speed, or visual memory. Our participants were, in general, very well-educated and mostly Caucasian. Additional studies are needed to determine if the results apply more broadly to those at the lower end of educational attainment and in more diverse ethnic and racial groups. Measuring practice effects at yearly intervals may not be practical in a clinical setting and therefore the generalizability of our findings may be limited. There is some evidence that assessment of practice effects within a shorter test-retest interval is sensitive to early stages of AD, including aMCI (Darby, Maruff, Collie, & McStephen, 2002; Duff et al., 2011), but additional studies are needed to determine whether a shorter test-retest interval will be sensitive to preclinical AD. We do not collect information on prior exposure to cognitive testing at baseline, therefore we cannot be sure that all participants were naïve at baseline to the experience of cognitive testing and to the content of individual measures. Finally, limiting our practice effects model to the first three assessments may underrepresent the influence of retesting. Although practice effects appear to be maximal on the first two retests (Calamia et al., 2012; Ivnik et al., 2000; Thorndike, 1922), it is likely that they continue to impact performance at the fourth assessment and beyond.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

References

- Aggleton JP, Brown MW. Episodic memory, amnesia, and the hippocampal-anterior thalamic axis. *The Behavioral and Brain Sciences*. 1999; 22(3):425–44. discussion 444–89. [PubMed: 11301518]
- Albert MS, DeKosky ST, Dickson D, Dubois B, Feldman HH, Fox NC, et al. The diagnosis of mild cognitive impairment due to Alzheimer's disease: recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. Presented at the Alzheimer's & dementia: the journal of the Alzheimer's Association. 2011; 7:270–279.10.1016/j.jalz.2011.03.008
- Armitage SG. An analysis of certain psychological tests used for the evaluation of brain injury. *Psychological Monographs*. 1946; 60(1):i–48.10.1037/h0093567
- Bateman RJ, Xiong C, Benzinger TLS, Fagan AM, Goate A, Fox NC, et al. Clinical and biomarker changes in dominantly inherited Alzheimer's disease. *The New England Journal of Medicine*. 2012; 367(9):795–804.10.1056/NEJMoa1202753 [PubMed: 22784036]
- Bates, DM. lme4: Mixed-effects modeling with R. 2010. <http://lme4.R-Forge.R-Project.Org/Book>
- Berg L, McKeel DW Jr, Miller JP, Storandt M, Rubin EH, Morris JC, et al. Clinicopathologic studies in cognitively healthy aging and Alzheimer disease: relation of histologic markers to dementia severity, age, sex, and apolipoprotein E genotype. *Archives of Neurology*. 1998; 55(3):326. [PubMed: 9520006]
- Buckner RL, Snyder AZ, Shannon BJ, LaRossa G, Sachs R, Fotenos AF, et al. Molecular, structural, and functional characterization of Alzheimer's disease: evidence for a relationship between default activity, amyloid, and memory. *The Journal of Neuroscience: the Official Journal of the Society for Neuroscience*. 2005; 25(34):7709–7717.10.1523/JNEUROSCI.2177-05.2005 [PubMed: 16120771]
- Calamia M, Markon K, Tranel D. Scoring Higher the Second Time Around: Meta-Analyses of Practice Effects in Neuropsychological Assessment. *The Clinical Neuropsychologist*. 2012; 26(4):543–570.10.1080/13854046.2012.680913 [PubMed: 22540222]

- Cooper DB, Lacritz LH, Weiner MF, Rosenberg RN, Cullum CM. Category fluency in mild cognitive impairment: reduced effect of practice in test-retest conditions. *Alzheimer Disease & Associated Disorders*. 2004; 18(3):120–122. [PubMed: 15494616]
- Darby D, Maruff P, Collie A, McStephen M. Mild cognitive impairment can be detected by multiple assessments in a single day. *Neurology*. 2002; 59(7):1042–1046.10.1212/WNL.59.7.1042 [PubMed: 12370459]
- Dickerson BC, Stoub TR, Shah RC, Sperling RA, Killiany RJ, Albert MS, et al. Alzheimer-signature MRI biomarker predicts AD dementia in cognitively normal adults. *Neurology*. 2011; 76(16): 1395–1402.10.1212/WNL.0b013e3182166e96 [PubMed: 21490323]
- Dodge HH, Wang CN, Chang CCH, Ganguli M. Terminal decline and practice effects in older adults without dementia The MoVIES project. *Neurology*. 2011; 77(8):722–730.10.1212/WNL.0b013e31822b0068 [PubMed: 21832224]
- Duff K, Beglinger LJ, Schultz SK, Moser DJ, McCaffrey RJ, Haase RF, et al. Practice effects in the prediction of long-term cognitive outcome in three patient samples: a novel prognostic index. *Archives of Clinical Neuropsychology: the Official Journal of the National Academy of Neuropsychologists*. 2007; 22(1):15–24.10.1016/j.acn.2006.08.013 [PubMed: 17142007]
- Duff K, Lyketsos CG, Beglinger LJ, Chelune G, Moser DJ, Arndt S, et al. Practice effects predict cognitive outcome in amnesic mild cognitive impairment. *The American Journal of Geriatric Psychiatry: Official Journal of the American Association for Geriatric Psychiatry*. 2011; 19(11): 932–939.10.1097/JGP.0b013e318209dd3a [PubMed: 22024617]
- Fine JP, Gray RJ. A proportional hazards model for the subdistribution of a competing risk. *Journal of the American Statistical Association*. 1999; 94(446):496–509.
- Galvin JE, Powlishta KK, Wilkins K, McKeel DW, Xiong C, Grant E, et al. Predictors of Preclinical Alzheimer Disease and Dementia A Clinicopathologic Study. *Archives of Neurology*. 2005; 62(5): 758–765.10.1001/archneur.62.5.758 [PubMed: 15883263]
- Goldberg TE, Keefe RSE, Goldman RS, Robinson DG, Harvey PD. Circumstances under which practice does not make perfect: a review of the practice effect literature in schizophrenia and its relevance to clinical treatment studies. *Neuropsychopharmacology: Official Publication of the American College of Neuropsychopharmacology*. 2010; 35(5):1053–1062.10.1038/npp.2009.211 [PubMed: 20090669]
- Goodglass H, Kaplan E. Boston diagnostic aphasia examination booklet. 1983
- Grober E, Buschke H, Crystal H, Bang S, Dresner R. Screening for dementia by memory testing. *Neurology*. 1988; 38(6):900–903. [PubMed: 3368071]
- Grober E, Hall CB, Lipton RB, Zonderman AB, Resnick SM, Kawas C. Memory impairment, executive dysfunction, and intellectual decline in preclinical Alzheimer's disease. *Journal of the International Neuropsychological Society*. 2008; 14(2):266–278.10.1017/S1355617708080302 [PubMed: 18282324]
- Hausknecht JP, Halpert JA, Di Paolo NT, Moriarty Gerrard MO. Retesting in selection: A meta-analysis of coaching and practice effects for tests of cognitive ability. *Journal of Applied Psychology*. 2007; 92(2):373–385.10.1037/0021-9010.92.2.373 [PubMed: 17371085]
- Howieson DB, Carlson NE, Moore MM, Wasserman D, Abendroth CD, Payne-Murphy J, Kaye JA. Trajectory of mild cognitive impairment onset. *Journal of the International Neuropsychological Society*. 2008; 14(2):192–198.10.1017/S1355617708080375 [PubMed: 18282317]
- Ivnik RJ, Smith GE, Petersen RC, Boeve BF, Kokmen E, Tangalos EG. Diagnostic accuracy of four approaches to interpreting neuropsychological test data. *Neuropsychology*. 2000; 14(2):163–177. [PubMed: 10791857]
- Jack CR Jr, Knopman DS, Jagust WJ, Shaw LM, Aisen PS, Weiner MW, et al. Hypothetical model of dynamic biomarkers of the Alzheimer's pathological cascade. *The Lancet Neurology*. 2010; 9(1): 119–128.10.1016/S1474-4422(09)70299-6 [PubMed: 20083042]
- Jack CR Jr, Knopman DS, Jagust WJ, Petersen RC, Weiner MW, Aisen PS, et al. Tracking pathophysiological processes in Alzheimer's disease: an updated hypothetical model of dynamic biomarkers. *The Lancet Neurology*. 2013; 12(2):207–216.10.1016/S1474-4422(12)70291-0 [PubMed: 23332364]

- Johnson DK, Storandt M, Morris JC, Galvin JE. Longitudinal study of the transition from healthy aging to Alzheimer disease. *Archives of Neurology*. 2009; 66(10):1254–1259.10.1001/archneurol.2009.158 [PubMed: 19822781]
- Johnson DK, Storandt M, Morris JC, Langford ZD, Galvin JE. Cognitive profiles in dementia: Alzheimer disease vs healthy brain aging. *Neurology*. 2008; 71(22):1783–1789.10.1212/01.wnl.0000335972.35970.70 [PubMed: 19029518]
- Knopman DS, Jack CR, Wiste HJ, Weigand SD, Vemuri P, Lowe V, et al. Short-term clinical outcomes for stages of NIA-AA preclinical Alzheimer disease. *Neurology*. 2012; 78(20):1576–1582.10.1212/WNL.0b013e3182563bbe [PubMed: 22551733]
- Lievens F, Buyse T, Sackett PR. Retest effects in operational selection settings: Development and test of a framework. *Personnel Psychology*. 2005; 58(4):981–1007.10.1111/j.1744-6570.2005.00713.x
- Lievens F, Reeve CL, Heggstad ED. An examination of psychometric bias due to retesting on cognitive ability tests in selection settings. *Journal of Applied Psychology*. 2007; 92(6):1672–1682.10.1037/0021-9010.92.6.1672 [PubMed: 18020804]
- Machulda MM, Pankratz VS, Christianson TJ, Ivnik RJ, Mielke MM, Roberts RO, et al. Practice Effects and Longitudinal Cognitive Change in Normal Aging vs. Incident Mild Cognitive Impairment and Dementia in The Mayo Clinic Study of Aging. *The Clinical Neuropsychologist*. 2013:1–18.10.1080/13854046.2013.836567 [PubMed: 23438260]
- McCaffrey, RJ.; Duff, K.; Westervelt, HJ. *Practitioner’s Guide to Evaluating Change With Intellectual Assessment Instruments*. Kluwer Academic Pub; 2000.
- Morris JC. The Clinical Dementia Rating (CDR): current version and scoring rules. *Neurology*. 1993; 43(11):2412–2414. [PubMed: 8232972]
- Petersen RC, Smith GE, Waring SC, Ivnik RJ, Tangalos EG, Kokmen E. Mild cognitive impairment: clinical characterization and outcome. *Archives of Neurology*. 1999; 56(3):303–308. [PubMed: 10190820]
- Price JL, Morris JC. Tangles and plaques in nondemented aging and “preclinical” Alzheimer’s disease. *Annals of Neurology*. 1999; 45(3):358–368. [PubMed: 10072051]
- Price JL, McKeel DW, Buckles VD, Roe CM, Xiong C, Grundman M, et al. Neuropathology of nondemented aging: presumptive evidence for preclinical Alzheimer disease. *Neurobiology of Aging*. 2009; 30(7):1026–1036.10.1016/j.neurobiolaging.2009.04.002 [PubMed: 19376612]
- Salthouse TA, Schroeder DH, Ferrer E. Estimating Retest Effects in Longitudinal Assessments of Cognitive Functioning in Adults Between 18 and 60 Years of Age. *Developmental Psychology*. 2004; 40(5):813–822.10.1037/0012-1649.40.5.813 [PubMed: 15355168]
- Sheline YI, Raichle ME, Snyder AZ, Morris JC, Head D, Wang S, Mintun MA. Amyloid plaques disrupt resting state default mode network connectivity in cognitively normal elderly. *Biological Psychiatry*. 2010; 67(6):584–587.10.1016/j.biopsych.2009.08.024 [PubMed: 19833321]
- Sperling RA, Aisen PS, Beckett LA, Bennett DA, Craft S, Fagan AM, et al. Toward defining the preclinical stages of Alzheimer’s disease: recommendations from the National Institute on Aging-Alzheimer’s Association workgroups on diagnostic guidelines for Alzheimer’s disease. Presented at the Alzheimer’s & dementia: the journal of the Alzheimer’s Association. 2011; 7:280–292.10.1016/j.jalz.2011.03.003
- Storandt M, Grant EA, Miller JP, Morris JC. Longitudinal course and neuropathologic outcomes in original vs revised MCI and in pre-MCI. *Neurology*. 2006; 67(3):467–473.10.1212/01.wnl.0000228231.26111.6e [PubMed: 16894109]
- Theisen ME, Rapport LJ, Axelrod BN, Brines DB. Effects of practice in repeated administrations of the Wechsler Memory Scale Revised in normal adults. *Assessment*. 1998; 5(1):85–92. [PubMed: 9458344]
- Thorndike EL. Practice Effects in Intelligence Tests. *Journal of Experimental Psychology*. 1922; 5(2): 101–107.10.1037/h0074568
- Tulving E. The effects of presentation and recall of material in free-recall learning. *Journal of Verbal Learning and Verbal Behavior*. 1967; 6(2):175–184.
- Tulving E. Episodic memory: From mind to brain. *Annual Review of Psychology*. 2002; 53(1):1–25.

- Vos SJ, Xiong C, Visser PJ, Jasiolec MS, Hassenstab J, Grant EA, et al. Preclinical Alzheimer's disease and its outcome: a longitudinal cohort study. *Lancet Neurology*. 2013; 12(10):957–965.10.1016/S1474-4422(13)70194-7 [PubMed: 24012374]
- Wechsler, D. WAIS Manual: Wechsler Adult Intelligence Scale. 1955.
- Wechsler, D. WMS-R: Wechsler Memory Scale-Revised: Manual. 1987.
- Wechsler, D.; Stone, CP. Wechsler memory scale manual. 1973.
- Zanetti O, Zanieri G, Giovanni GD, De Vreese LP, Pezzini A, Metitieri T, Trabucchi M. Effectiveness of procedural memory stimulation in mild Alzheimer's disease patients: A controlled study. *Neuropsychological Rehabilitation*. 2001; 11(3–4):263–272.10.1080/09602010042000088

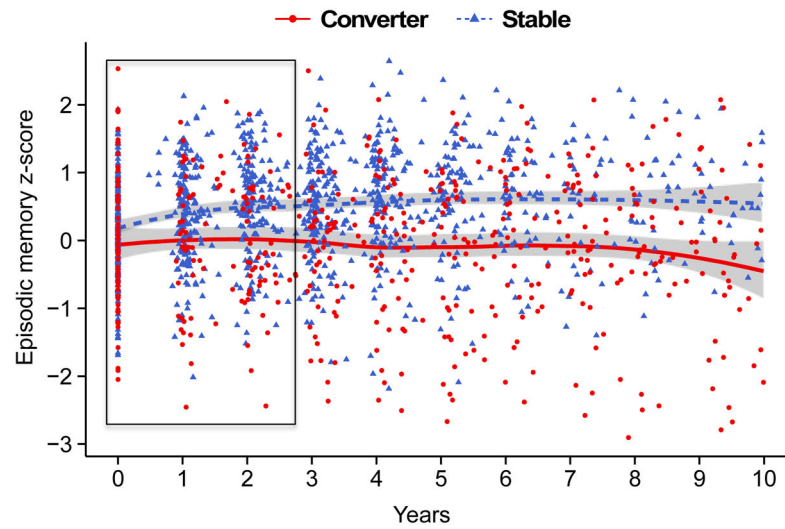


Figure 1. Plot of practice effects with LOESS curves for the episodic memory composite z-score in cognitively normal (CDR 0, n=263) participants who remained normal (“Stable”, n=197), or progressed to dementia (“Progressors”, CDR>0, n=66) after a minimum of 4 assessments. Practice effects were defined as the ordinary least squares slope of cognitive performance across the first three annual assessments, as indicated by the area within the rectangle.

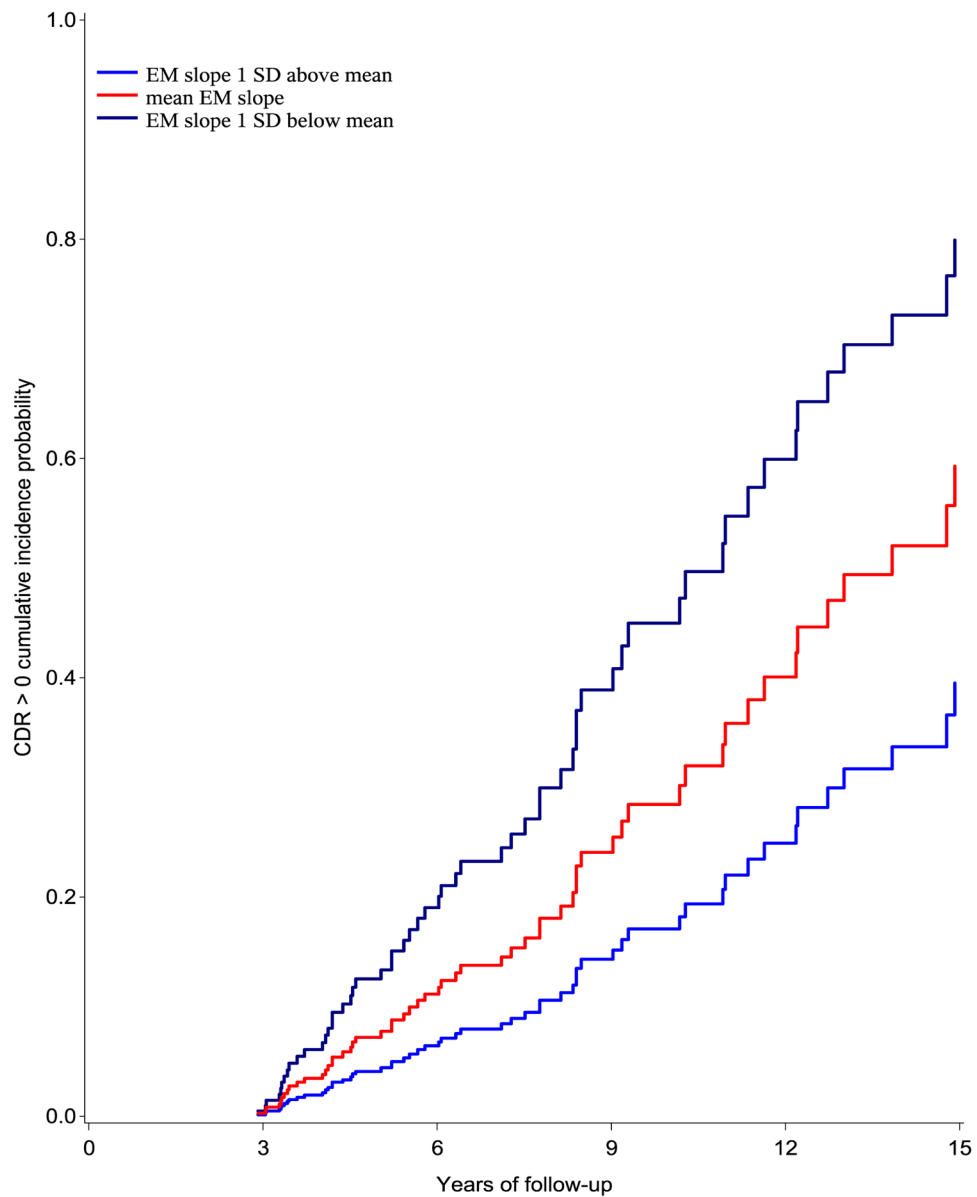


Figure 2. Cumulative incidence probability of progressing to a CDR=0.5 with an AD diagnosis. The ordinary least squares slope of episodic memory composite performance across the first three visits was used to predict survival. Curves were adjusted for age, gender, education, and APOE genotype. EM = Episodic Memory composite. Participants showing clear evidence of practice effects (in blue) had the lowest risk of progression.

Table 1

Demographic and Cognitive Performance at Baseline.

Demographics	Stable			Progressors			
	n	Mean	SD	Range	Mean	SD	Range
Female (%)	197						66 (34.55%)
Caucasian (%)	119 (60.41%)						59 (89.40%)
APOE4 (%)	178 (90.36%)						26 (39.40%)
Mortality (%) ^{***}	55 (28.21%)						14 (22.21%)
	11 (5.58%)						
<hr/>							
	Mean	SD	Range	Mean	SD	Range	
Age at Entry (years) ^{***}	71.74	5.60	65.05 - 88.81	74.67	6.64	65.07 - 88.13	
Education (years)	15.50	2.79	8 - 23	15.30	3.41	8 - 29	
Follow Ups (n assessments) ^{***}	6.63	3.86	4 - 27	10.06	4.11	4 - 23	
Years to Diagnosis	--	--	--	7.24	3.76	2.92 - 20.70	
Mini Mental Status Exam	29.02	1.13	24 - 30	29.00	1.17	26 - 30	
CDR Sum of Boxes ^a	0.04	0.14	0 - 0.5	0.08	0.22	0 - 1	
<hr/>							
Cognitive Scores (Baseline)	Mean	SD	Range	Mean	SD	Range	
Global Cognition Composite [*]	0.21	0.51	-1.72 - 1.65	-0.01	0.60	-1.29 - 1.20	
Episodic Memory Composite	0.18	0.71	-1.69 - 1.93	-0.08	0.99	-2.05 - 2.53	
WMS-R Logical Memory Immediate ^{**}	13.54	3.63	4 - 21	9.73	4.38	2 - 17	
WMS-R Logical Memory Delayed ^b	12.43	3.90	3 - 20	9.09	4.61	1 - 16	
FCSRT Free Recall	29.55	5.08	15 - 43	26.10	6.82	16 - 42	
WMS Associate Learning ^b	14.03	3.37	5.5 - 21	13.23	3.69	6.5 - 20	
Visuospatial Composite [*]	0.37	0.76	-1.88 - 2.97	0.02	0.67	-1.13 - 1.49	
WAIS Block Design	32.39	8.31	6 - 47	30.65	7.85	14 - 47	
WAIS Digit Symbol Substitution	49.54	10.62	12 - 75	45.67	11.09	2 - 72	
Trailmaking Test Part A [*]	32.72	10.38	15 - 88	37.66	12.48	17 - 70	
Trailmaking Test Part B	83.91	32.83	28 - 180	96.49	31.48	48 - 180	
Semantic Composite	0.16	0.71	-2.59 - 1.64	-0.05	0.83	-2.27 - 1.45	

Demographics	Stable			Progressors				
WAIS Information	21.71	3.93	7	29	20.92	5.02	6	28
Boston Naming Test	55.42	5.63	18	60	53.94	5.33	38	60
Animal Naming	20.54	5.07	9	34	19.33	5.75	9	31
Executive Function Composite	0.12	0.68	-1.84	1.46	0.04	0.70	-1.57	1.48
WMS Mental Control	7.57	1.72	2	9	7.34	1.91	2	9
WMS Digit-Span Forward	6.63	1.22	4	8	6.55	1.21	4	8
WMS Digit-Span Backward	4.81	1.22	2	8	4.63	1.38	1	7
Letter Fluency (S & P)	31.01	9.87	-1.87	1.46	31.05	10.02	-1.57	1.48

Abbreviations: APOE4 = one or more copy of apolipoprotein ε4 allele, CDR = Clinical Dementia Rating scale.

Note: comparisons of cognitive scores adjusted for age, gender, and education;

^aCDR Sum of Boxes higher score = worse performance. Samples sizes may vary slightly due to missing data.

^b p < .10

* p < 0.05;

** p < 0.01;

*** p < 0.001

Table 2

Cognitive Performance Across First Three Visits

	Baseline		Visit 2		Visit 3	
	Mean	SD	Mean	SD	Mean	SD
Stable (n=197)						
Global Cognition Composite (z-score)	0.21	0.51	0.28	0.50	0.34	0.55
Episodic Memory Composite (z-score)	0.18	0.71	0.35	0.72	0.51	0.73
WMS-R Logical Memory Immediate	13.54	3.63	13.87	3.60	14.60	3.53
WMS-R Logical Memory Delayed	12.43	3.90	13.03	3.61	13.86	3.63
FCSRT Free Recall	29.55	5.08	30.74	5.40	31.49	5.40
WMS Associate Learning	14.03	3.37	14.76	3.19	15.54	7.00
Visuospatial Composite (z-score)	0.37	0.76	0.39	0.75	0.44	0.79
WAIS Block Design	32.39	8.31	32.49	8.38	33.12	7.77
WAIS Digit Symbol Substitution	49.54	10.62	50.79	10.45	51.25	10.64
Trailmaking Test Part A	32.72	10.38	34.93	9.77	33.25	11.74
Trailmaking Test Part B	83.91	32.83	81.57	32.04	82.47	33.86
Semantic Composite (z-score)	0.16	0.71	0.22	0.66	0.22	0.71
WAIS Information	21.71	3.93	22.01	3.74	22.38	3.82
Boston Naming Test	55.42	5.63	55.68	4.64	55.59	4.88
Animal Naming	20.54	5.07	21.02	4.92	20.62	5.60
Executive Function Composite (z-score)	0.12	0.68	0.15	0.67	0.16	0.72
WMS Mental Control	7.57	1.72	7.56	1.68	7.47	1.74
WMS Digit Span Forward	6.63	1.22	6.72	1.09	6.68	1.23
WMS Digit Span Backward	4.81	1.22	4.96	1.23	4.89	1.28
Letter Fluency (S & P)	31.01	9.87	30.55	10.11	31.68	10.05
Progressors (n=66)						
Global Cognition Composite (z-score)	-0.01	0.60	0.08	0.58	0.04	0.63
Episodic Memory Composite (z-score)	-0.08	0.99	0.07	0.97	0.01	1.00
WMS-R Logical Memory Immediate	9.73	4.38	11.47	4.23	11.15	3.39
WMS-R Logical Memory Delayed	9.09	4.61	10.06	5.91	9.25	4.44
FCSRT Free Recall	26.10	6.82	27.20	6.38	26.67	7.17

Stable (n=197)	Baseline		Visit 2		Visit 3	
	Mean	SD	Mean	SD	Mean	SD
WMS Associate Learning	13.23	3.69	13.89	4.25	13.81	3.93
Visuospatial Composite (z-score)	0.02	0.67	0.10	0.71	0.04	0.79
WAIS Block Design	30.65	7.85	31.33	7.28	30.11	7.33
WAIS Digit Symbol Substitution	45.67	11.09	45.32	11.18	45.42	10.93
Trailmaking Test Part A	37.66	12.48	33.55	11.45	37.37	12.84
Trailmaking Test Part B	96.49	31.48	92.31	28.52	93.66	34.87
Semantic Composite (z-score)	-0.05	0.83	-0.06	0.75	-0.07	0.86
WAIS Information	20.92	5.02	20.63	4.57	20.72	5.23
Boston Naming Test	53.94	5.33	54.20	4.75	53.88	5.63
Animal Naming	19.33	5.75	19.53	5.40	19.39	5.65
Executive Function Composite (z-score)	0.04	0.70	0.18	0.66	0.15	0.70
WMS Mental Control	7.34	1.91	7.52	1.87	7.60	1.52
WMS Digit Span Forward	6.55	1.21	6.65	1.18	6.65	1.19
WMS Digit Span Backward	4.63	1.38	5.03	1.18	4.86	1.38
Letter Fluency (S & P)	31.05	10.02	31.98	10.45	31.82	10.79

Abbreviations: WMS-R = Wechsler Memory Scale-Revised, FCSRT = Free and Cued Selective Reminding Test, WMS = Wechsler Memory Scale, WAIS = Wechsler Adult Intelligence Scale. *Note: Scores presented are raw scores unless indicated otherwise. Sample sizes may vary slightly due to missing data.*

Table 3

Slopes of cognitive performance presented as group (Stable (n=197) and Progressors (n=66)) by time (years from baseline assessment) interactions. The estimate term represents the slope of the Progressor group compared to the Stable group as the reference. A) Slopes for first three visits when practice effects are most pronounced. B) Slopes across all visits. Covariates: age at baseline, gender, education, cognitive score at baseline, and age at baseline by time interaction.

Cognitive Composites	Estimate	SE	t	p
A) First Three Visits Only				
Visuospatial Composite	0.012	0.031	0.380	0.707
Semantic Composite	-0.027	0.030	-0.920	0.356
Executive Function Composite	0.046	0.032	1.420	0.158
Episodic Memory Composite	-0.107	0.041	-2.620	0.009
Global Cognition Composite	-0.015	0.019	-0.810	0.420
B) All Visits				
Visuospatial Composite	-0.073	0.010	6.984	<0.001
Semantic Composite	-0.082	0.007	6.625	<0.001
Executive Function Composite	-0.044	0.008	5.525	<0.001
Episodic Memory Composite	-0.160	0.015	10.240	<0.001
Global Cognition Composite	-0.088	0.009	10.010	<0.001

Abbreviations: WMS = Wechsler Memory Scale, SE = standard error. *Note: Sample sizes may vary slightly due to missing data.*

Changepoint models examining the effect of practice on cognitive trajectories in all participants on the episodic memory composite and the global cognitive composite. Models with a single linear slope assume no effect of practice on rate of decline. Models with a changepoint at visit 3 assume a separate slope across the first 3 visits.

Table 4

Composites	Model with Changepoint at Visit 3				Single Linear Slope			
	Effect	Estimate	SE	p	Effect	Estimate	SE	p
Episodic Memory Composite (n=263)	Slope of Visits 1 to 3	0.117	0.017	<.001	Slope Across All Visits	0.005	0.011	0.66
	Slope of Visits >3	-0.064	0.013	<.001				
Global Cognitive Composite (n=233)	Slope of Visits 1 to 3	0.046	0.008	<.001	Slope Across All Visits	-0.014	0.006	0.02
	Slope of Visits >3	-0.048	0.008	<.001				